Description

In Vivo Coding has also been labeled “literal coding,” “verbatim coding,” “inductive coding,” “indigenous coding,” “natural coding,” and “emic coding” in selected methods literature. In this manual, In Vivo Coding will be used since it is the most well-known label.

The root meaning of in vivo is “in that which is alive,” and as a code refers to a word or short phrase from the actual language found in the qualitative data record, “the terms used by [participants] themselves” (Strauss, 1987, p. 33).

Folk or *indigenous* terms are participant-generated words from members of a particular culture, subculture, microculture, or counterculture. Folk terms indicate the existence of the group’s cultural categories (McCurdy, Spradley, & Shandy, 2005, p. 26). For example, some homeless youth say that they “sp’ange” (ask passers-by for “spare change”). Digital culture has created such terms and acronyms as “tweet,” “LOL,” and “FML.” In Vivo Coding a subculture’s unique vocabulary or *argot* is one method of extracting these indigenous terms (and see Domain and Taxonomic Coding for more specific categorization guidelines).

Applications

In Vivo Coding is appropriate for virtually all qualitative studies, but particularly for beginning qualitative researchers learning how to code data, and studies that prioritize and honor the participant's voice. In Vivo Coding is one of the methods to employ during grounded theory's Initial Coding but can be used with several other coding methods in this manual.

In Vivo Coding is particularly useful in educational ethnographies with youth. The child and adolescent voices are often marginalized, and coding with their actual words enhances and deepens an adult's understanding of their discourses, cultures, and worldviews. In Vivo Coding is also quite applicable to action, participatory, and practitioner research (Coghlan & Brannick, 2014; Fox, Martin, & Green, 2007; Stringer, 2014) since one of the genre's primary goals is to adhere to the "*verbatim principle*, using terms and concepts drawn from the words of the participants themselves. By doing so [researchers] are more likely to capture the meanings inherent in people's experience" (Stringer, 2014, p. 140).

After decades of coding qualitative data, I have personally found In Vivo Coding to be my first "go-to" method with interview transcript data, regardless of the study's research questions or methodological approach.

Example

An adult female interviewer talks to Tiffany, a 16-year-old teenage girl, about her friendships at high school. Note how all In Vivo Codes are placed in quotation marks, and how virtually each line of data gets its own code. This is coding as a “splitter”:

Code example 6.2

I [1] hated school last year. [2] Freshman year, it was awful, I hated it. And [3] this year’s a lot better actually. Um, I [4] don’t know why. I guess, over the summer I kind of [5] stopped caring about what other people thought and cared more about, just, I don’t know. It’s [6] hard to explain. I [7] found stuff out about myself, and so I went back, and all of a sudden I found out that when I [8] wasn’t trying so hard to [9] have people like me and to do [10] what other people wanted, people [11] liked me more. It was [12] kind of strange. Instead of [13] trying to please them all the time, they liked me more when I [14] wasn’t trying as hard. And, I don’t know, like every-, everybody might, um, people who are just, kind of, [15] friends got closer to me. And people who didn’t really know me [16] tried to get to know me. [17] I don’t know.	[1] “hated school” [2] “freshman year awful” [3] “this year’s better” [4] “don’t know why” [5] “stopped caring” [6] “hard to explain” [7] “found stuff out”

	[8] “wasn’t trying so hard” [9] “have people like me” [10] “what other people wanted” [11] “liked me more” [12] “kind of strange” [13] “trying to please them” [14] “wasn’t trying as hard” [15] “friends got closer” [16] “tried to know me” [17] “i don’t know”

Key writers of grounded theory advocate meticulous work and that an In Vivo (or other) Code should appear next to every line of data. Depending on your goals, In Vivo Codes can be applied with less frequency, such as one word or phrase for every three to five sentences. In the interview excerpt above, rather than 17 In Vivo Codes I could have limited the number to 4—coding as a “lumper”:

Code example 6.3

I hated school last year. [1] Freshman year, it was awful, I hated it. And this year’s a lot better actually. Um, I don’t know why. I guess, over the summer I kind of stopped caring about what other people thought and cared more about, just, I don’t know. It’s hard to explain. I [2] found stuff out about myself, and so I went back, and all of a sudden I found out that when I [3] wasn’t trying so hard to have people like me and to do what other people wanted, people liked me more. It was kind of strange. Instead of trying to please them all the time, they liked me more when I wasn’t trying as hard. And, I don’t know, like every-, everybody might, um, people who are just, kind of, [4] friends got closer to me. And people who didn’t really know me tried to get to know me. I don’t know.	[1] “freshman year awful” [2] “found stuff out” [3] “wasn’t trying so hard” [4] “friends got closer”

Analysis

As you read interview transcripts or other documents that feature participant voices, attune yourself to words and phrases that seem to call for bolding, underlining, italicizing, highlighting, or vocal emphasis if spoken aloud. Their salience may be attributed to such features as impacting nouns, action-oriented verbs, evocative vocabulary, clever or ironic phrases, similes and metaphors, etc. If the same words, phrases, or variations thereof are used often by the participant (such as “I don’t know” in the example above), and seem to merit an In Vivo Code, apply it. In Vivo Codes “can provide a crucial check on whether you have grasped what is significant” to the participant, and may help crystallize and condense meanings (Charmaz, 2014, p. 135). Thus, keep track of codes that are participant-inspired rather than researcher-generated by always putting In Vivo Codes in quotation marks: “hated school”.

There is no fixed rule or formula for an average number of codes per page or a recommended ratio of codes to text. Trust your instincts with In Vivo Coding. When something in the data appears to stand out, apply it as a code. Researcher reflection through analytic memo writing, coupled with second cycle coding, will condense the number of In Vivo Codes and provide a reanalysis of your initial work. Strauss (1987, p. 160) also recommends that researchers examine In Vivo Codes not just as themes but as possible *dimensions* of categories—that is, the continuum or range of a property.

An initial analytic tactic with codes is to eyeball (i.e., analytically browse) the codes list to inventory the contents and to discern any particular patterns. Extracting and reorganizing the codes list in alphabetical order through Word or Excel’s “Data Sort” function may enable you to detect some similar clusters. For example, the alphabetized array reads:

“DON’T KNOW WHY”

“FOUND STUFF OUT”

“FRESHMAN YEAR AWFUL”

“FRIENDS GOT CLOSER”

“HARD TO EXPLAIN”

“HATED SCHOOL”

“HAVE PEOPLE LIKE ME”

“I DON’T KNOW”

“KIND OF STRANGE”

“LIKED ME MORE”

“STOPPED CARING”

“THIS YEAR’S BETTER”

“TRIED TO KNOW ME”

“TRYING TO PLEASE THEM”

“WASN’T TRYING AS HARD”

“WASN’T TRYING SO HARD”

“WHAT OTHER PEOPLE WANTED

The alphabetical proximity of “trying to please them”, “wasn’t trying as hard”, and “wasn’t trying so hard” suggests a possible category or theme at work, prompting a marginal jotting or analytic memo on “trying”.

A second method for initially organizing the array of In Vivo Codes is to list them on a text editing page and then cut and paste them into outlined clusters that suggest categories of belonging and an order of some kind (such as hierarchical, chronological, mirco to macro, etc. See Chapter 4 for another extended example of In Vivo Coding). Using the 17 In Vivo Codes from the split-coding example above, the outline reads as follows:

I. **“HATED SCHOOL”**
    A. “FRESHMAN YEAR AWFUL”
II. **“STOPPED CARING”**
    A. “WHAT OTHER PEOPLE WANTED”
        1. “HAVE PEOPLE LIKE ME”
        2. “TRYING TO PLEASE THEM”
    B. “FOUND STUFF OUT”
        1. “WASN’T TRYING SO HARD”
        2. “WASN’T TRYING AS HARD”
III. **“THIS YEAR’S BETTER”**
    A. “FRIENDS GOT CLOSER”
    B. “LIKED ME MORE”
    C. “TRIED TO KNOW ME”
IV. **“DON’T KNOW WHY”**
    A. “I DON’T KNOW”
    B. “KIND OF STRANGE”
    C. “HARD TO EXPLAIN”

Remember that memos are a critical component of grounded theory’s coding processes, and *memo writing also serves as a code-, category-, theme-, and concept-generating method*. An analytic memo excerpt based on the coding example above reads as follows (and note that In Vivo Codes and participant quotes are included throughout):

---

25 May 2011

CODE: “I DON’T KNOW”

Tiffany is genuinely puzzled (“DON’T KNOW WHY”, “HARD TO EXPLAIN”) by a paradox of sustaining quality friendships: “when I wasn’t trying so hard to have people like me and to do what other people wanted, people liked me more.” At age 16, she is learning “about myself”—who she is and wants to become, rather than what others want or expect from her. Just as there is documented and developmental *emotional ambivalence* in middle childhood, perhaps adolescence has its equivalent stage called SOCIAL AMBIVALENCE, concurrent with the individual’s IDENTITY WORK.

In Vivo Codes capture “behaviors or processes which will explain to the analyst how the basic problem of the actors is resolved or processed” (Strauss, 1987, p. 33) and help preserve participants’ meanings of their views and actions in the coding itself (Charmaz, 2014). In Vivo Codes can also provide imagery, symbols, and metaphors for rich category, theme, and concept development, plus evocative content for arts-based interpretations of the data. Using some of Tiffany’s own language, a poetic reconstruction (called “found poetry” or “poetic transcription”) of the above vignette’s codes and transcript excerpts might read:

Freshman year:

awful,

hated school. …

Over the summer:

stopped caring about what others thought,

found stuff out about myself. …

This year's better:

friends got closer,

tried to know me,

liked me more. …

Don't know why:

kind of strange,

hard to explain. …

This year's better. (Saldaña, 2011b, p. 129)

Playwright and verbatim theatre performer Anna Deavere Smith asserts that people speak in "organic poems" through everyday discourse. Thus, In Vivo Coding is one strategy for getting at the organic poetry inherent in a participant.

In Vivo Codes could be used as the sole coding method for the first cycle of data analysis, and the sole method of choice for small-scale studies, but that may limit the researcher's perspective on the data, a perspective that can contribute to more conceptual and theoretical views about the phenomenon or process. Sometimes the participant says it best; sometimes the researcher does. Be prepared and willing to mix and match coding methods as you proceed with data analysis.

Several CAQDAS programs make In Vivo Coding easy by permitting the analyst to select a word or short phrase from the data, clicking a dedicated icon, and assigning the selected text *as* an In Vivo Code. But be aware that some CAQDAS functions will retrieve multiple text units only if they share the *exact* same code you have applied to them.

In Vivo Coded data most often are so unique that they will require careful review and self-categorization into an NVivo node, for example. Also, selected CAQDAS programs may not permit the use of quotation marks to accompany and indicate an In Vivo Code entry. Thus, find an alternative format for the code (e.g., all CAPS or a dedicated color font) in lieu of quotation marks, if necessary.

Some recommended ways to further analyze In Vivo Codes are (see Appendix B):

- second cycle coding methods
- action and practitioner research (Altrichter, Posch, & Somekh, 1993; Coghlan & Brannick, 2014; Fox et al., 2007; Stringer, 2014)
- case studies (Merriam, 1998; Stake, 1995)
- discourse analysis (Bischoping & Gazso, 2016; Gee, 2011; Rapley, 2018; Willig, 2015)
- domain and taxonomic analysis (Schensul et al., 1999b; Spradley, 1979, 1980)
- frequency counts (LeCompte & Schensul, 2013)
- grounded theory (Bryant, 2017; Bryant & Charmaz, 2007, 2019; Charmaz, 2014; Corbin & Strauss, 2015; Stern & Porr, 2011)
- interactive qualitative analysis (Northcutt & McCoy, 2004)
- memo writing about the codes/themes (Charmaz, 2014; Corbin & Strauss, 2015; Glaser, 1978; Glaser & Strauss, 1967; Strauss, 1987)
- metaphoric analysis (Coffey & Atkinson, 1996; Lakoff & Johnson, 2003; Todd & Harrison, 2008)
- narrative inquiry and analysis (Bischoping & Gazso, 2016;

Clandinin & Connelly, 2000; Coffey & Atkinson, 1996; Cortazzi, 1993; Coulter & Smith, 2009; Daiute & Lightfoot, 2004; Holstein & Gubrium, 2012; Murray, 2015; Riessman, 2008)

- phenomenology (Giorgi & Giorgi, 2003; Smith, Flowers, & Larkin, 2009; Vagle, 2018; van Manen, 1990; Wertz et al., 2011)
- poetic and dramatic writing (Denzin, 1997, 2003; Glesne, 2011; Knowles & Cole, 2008; Leavy, 2015; Saldaña, 2005a, 2011a)
- polyvocal analysis (Hatch, 2002)
- portraiture (Lawrence-Lightfoot & Davis, 1997)
- qualitative evaluation research (Patton, 2008, 2015)
- sentiment analysis (Ignatow & Mihalcea, 2017; Liu, 2015)
- social media analysis (Kozinets, 2020; Paulus & Wise, 2019; Rogers, 2019; Salmons, 2016)
- thematic analysis (Auerbach & Silverstein, 2003; Boyatzis, 1998; Smith & Osborn, 2015).

Notes

Researchers new to coding qualitative data often find In Vivo Coding a safe and secure method with which to begin. But be wary of overdependence on the strategy because it can limit your ability to transcend to more conceptual and theoretical levels of analysis and insight.

For an example of how In Vivo Coding can be used with a strategic focus, see Metaphor Coding in Chapter 8. Also see Carolyn Lunsford Mears' (2009) outstanding book on interviewing and transcript analysis, *Interviewing for Education and Social Science Research*, which extracts and arranges the essentialized verbatim texts of participants into poetic mosaics.